



IMF Working Paper

European Department

The Negative Mean Output Gap

Prepared by Shekhar Aiyar and Simon Voigts¹

August 2019

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

Abstract

We argue that in an economy with downward nominal wage rigidity, the output gap is negative on average. Because it is more difficult to cut wages than to increase them, firms reduce employment more during downturns than they increase employment during expansions. This is demonstrated in a simple New Keynesian model with asymmetric wage adjustment costs. Using the model's output gap as a benchmark, we further show that common output gap estimation methods exhibit a systematic bias because they assume a zero mean. The bias is especially large in deep recessions when potential output tends to be most severely underestimated.

JEL Classification Numbers: E32, E61

Keywords: Output gap estimation, business cycles, fiscal policy, monetary policy

Author's E-Mail Address: SAiyar@imf.org; SVoigts@imf.org

¹ We thank Larry Christiano, Helge Berger, Gita Gopinath, Enrica Detragiache, Alfredo Cuevas, Bernd Droge, Jeromin Zettelmeyer and Bertrand Gruss for valuable comments, and Xiaobo Shao for excellent research assistance.

Contents

1	Introduction	3
1.1	Motivation and Contributions	3
1.2	Output Gaps, Wage Rigidity and Filters	3
2	Model	5
2.1	Summary of the model	5
2.2	Output gap definition	8
2.3	Calibration	8
3	The model's adjustment to shocks	9
3.1	Demand shocks	9
3.2	Supply shocks	10
3.3	Relation of the model output gap to the Friedman-Phelps 'natural rate'	11
4	Mean output gap in the model	12
5	Accuracy of output gap filters	15
5.1	The Hodrick-Prescott Filter	16
5.2	The IMF's multivariate filter	17
5.3	The zero-mean property of filters and shock identification	18
6	Conclusion	19

1 Introduction

1.1 Motivation and Contributions

The output gap is a central feature of modern macroeconomics. Positive and negative deviations from potential output define the amount of slack in an economy, provide the environment within which wage bargaining and price-setting occur, and correspond to booms and downturns in the business cycle. The size and evolution of the output gap is at the heart of both monetary and fiscal policy. Rules-based methodologies for setting interest rates – such as the Taylor Rule – operate within an output gap framework, while the minutes of any advanced economy monetary policy committee bear ample testimony to its role in policy deliberations. Likewise, fiscal policy is usually premised on some notion of the structural budget balance, which requires an estimate of the output gap.

This paper’s first contribution is to make and elaborate a basic claim about the output gap, namely that, properly defined, it should have a negative mean over the long run. A negative mean output gap is featured by some existing models such as Benigno and Ricci (2011), Dupraz, Nakamura and Steinsson (2017) and Schmitt-Grohé and Uribe (2016). However, to the best of our knowledge and despite its obvious policy relevance, no paper has explicitly focused on the output gap and demonstrated that the gap in a New-Keynesian model – which is typically formulated in terms of distance from the fully flexible price allocation – is equivalent to the definition relevant to policymakers, which has to do with the buildup of inflationary or deflationary pressures and has an intellectual pedigree reaching back to Friedman and Phelps. We make this link, and then argue that a negative mean output gap follows in a very intuitive way from downward nominal wage rigidity – a feature of the world that has overwhelming empirical support.

Our second contribution pertains to the estimation of the output gap. We show that standard output gap estimation methods (‘filters’ from now on) embed a symmetric view of fluctuations and constrain the estimated series average to zero. This leads to systematic bias in an economy where the true output gap is non-zero. Exploiting the fact that the true output gap is observable in a model economy, we apply standard filters to artificial data obtained from a simulation of the Abbritti and Fahr (2013) model, demonstrating that potential output is underestimated relative to its true (model-implied) level.

1.2 Output Gaps, Wage Rigidity and Filters

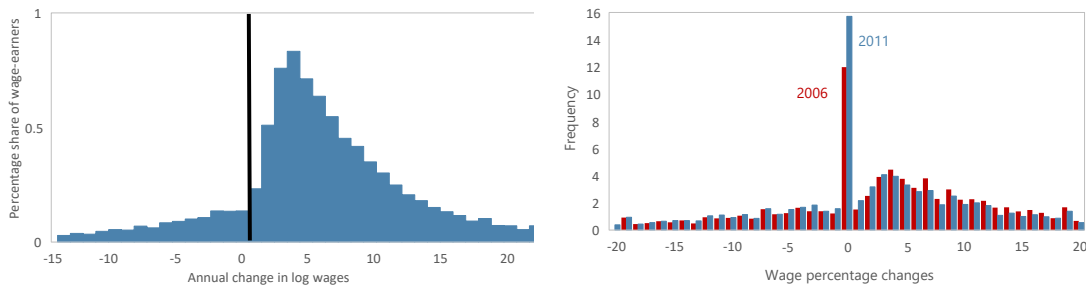
The output gap has a long intellectual history (see Congdon, 2008, for an excellent survey). The seminal work of Phillips (1958) documented an inverse relationship between the rate of change of wages and the unemployment rate, while Okun (1962) further defined deviations from potential output in terms of the gap between actual unemployment and ‘full employment’ (taken by Okun to be 4 percent). But the modern formulation of the output gap dates back to Friedman (1968) and Phelps (1967), who posited that the rate of wage increase was stable at only a single rate of unemployment, termed the ‘natural rate’. This natural rate of unemployment is isomorphic to a zero output gap, where the rate of inflation is stable. If output (and employment) rises above this level, inflation accelerates; if it falls below this level, inflation declines.

In New Keynesian models (see, for example Clarida, Gali and Gertler, 1999), it is price rigidities that cause deviations of output from its natural level. The fully flexible price allocation corresponds to a closed output gap, at which point there is no tendency for inflation to either rise or fall. We show that when a New Keynesian model is extended by labor search and matching frictions (e.g. Pissarides, 2000), the output gap corresponds closely to the concept of Friedman and Phelps. A positive demand shock coupled with a failure of prices and wages to adjust instantaneously moves the economy above potential, wages increase over several periods, the unemployment rate drops below its structural level, and inflation accelerates. A negative demand shock, on the other hand, causes unemployment to rise above its structural level, and the rate of inflation declines.

Our central argument is that the asymmetry introduced by downward (but not upward) nominal wage rigidity shifts the mean of the output gap below zero. There is overwhelming empirical evidence for the proposition that nominal wage cuts are rare, making the distribution of nominal wage changes

extremely non-normal. The International Wage Flexibility Project (IWFP)—a consortium of 40 researchers examining wage data in 16 countries—concludes that there is strong evidence of downward nominal wage rigidity, with some heterogeneity across countries (Dickens et al., 2007). Not only does a histogram of wage changes show very little weight below zero, but there is a sharp spike at zero, indicating a nominal bound beneath which wages cannot descend despite the ‘demand’ for lower real wages. Moreover, the extent to which nominal wage rigidity binds appears to become more acute in recessions. Daly et al. (2013) show that in the US, the share of workers whose wage changes were ‘frozen’ at zero jumped from 12% in 2006 to 16% in 2011, during the Great Recession. Notably, even during the very high unemployment environment of the Great Recession, very few workers experienced wage cuts, with the numbers edging up only slightly from 2006.

Figure 1: Evidence for downward nominal wage rigidity



Sources: Dickens et al. (2007) and Daly et al. (2013) (left and right chart respectively).

Downward nominal wage rigidity results in different dynamics for negative deviations from potential output relative to positive deviations. The intuition is simple. Reducing the nominal wage is difficult or impossible for employers. So the response to a negative shock requires relatively greater adjustment of quantities (employment) to compensate for the relatively lesser adjustment of prices (wages). Assuming that positive and negative shocks to the economy are symmetrically distributed, it follows that cumulative negative deviations from potential output are greater than cumulative positive deviations. Symmetric shocks result in an asymmetric, negative-mean cycle.

We formalize this intuition by studying the output gap in a simple New Keynesian business cycle model with downward nominal wage rigidity, taken from Abbritti and Fahr (2013). To preview our main result, the mean output gap is negative, amounting to about -1% of potential output. The negative mean is driven mainly by periods of elevated volatility, so the average is closer to zero when turbulent periods are excluded. The intuition is that downward nominal wage rigidity plays a more powerful role in deep recessions than in shallow downturns. If the downturn is shallow then real wages can adjust downwards even in the absence of nominal downward wage adjustments (for example, this would be achieved by constant nominal wages coupled with some positive rate of general inflation). In a deep recession, where general price inflation falls steeply (or even becomes negative), downward nominal wage rigidity is more binding, forcing greater adjustment in quantities. Thus the asymmetry introduced by downward nominal wage rigidities is particularly pronounced for large shocks.

Asymmetric shock adjustment, however, cannot be captured by standard output gap filters. We focus on two examples, showing that the cyclical component derived from an HP filter always has a mean of zero, while the same holds approximately for the IMF’s multivariate filter (Blagrove et al., 2015). To study how this causes estimation bias in an asymmetric world, we simulate the Abbritti and Fahr (2013) model to generate synthetic time series for macroeconomic variables, including the output gap. Since the true output gap in the model economy is observable, it can be used as a benchmark to assess the accuracy of the filters. We apply the filters to simulated data and compare the resulting estimates with the true, model-implied output gap. We find that the filters significantly underestimate potential output; in other words, they overestimate – in the positive direction – the output gap. For example, the bias of the HP filter amounts to about 1% of potential when we average over the full generated sample, and increases to about 3% when we focus on the most turbulent quarter of the sample. The use of model-generated data as the benchmark against which to assess the reliability of

output gap filters differentiates our paper from existing studies, which typically interpret final estimates as the benchmark for the true output gap.¹

The bias is explained by the incommensurability of asymmetry in the model and symmetry in the filters. In the model downward nominal wage rigidity causes cumulative negative deviations of output from potential to be greater than cumulative positive deviations, implying a negative mean output gap. Filter-estimated output gaps, by contrast, have a mean of zero, so negative and positive cumulative deviations of output from estimated potential must cancel out. If the true data generating process is indeed characterized by greater negative deviations, the only way that the estimation technique can produce a zero mean output gap is by ‘shifting down’ potential output, so that the estimated negative deviations remain in balance with the estimated positive deviations.

Our results are consistent with, and serve to illuminate, recent critiques of business cycle empirics. Coibion, Gorodnichenko and Ulate (2019) document the tendency of official estimates of potential output to move in line with cyclical shocks that have little long-run impact on the economy. They find that the way in which estimates of potential GDP respond to shocks can be well represented by fitting an HP filter to real time GDP shocks. Our work shows that fitting an HP filter to a time series characterized by underlying asymmetry generates estimates of potential output that (incorrectly) track demand shocks, and are therefore biased. Moreover, we show that the bias is most prominent in deep recessions. This could help explain the widespread downward revisions to potential output in the wake of Great Recession, a phenomenon noted by several authors (see e.g. Ball, 2014 or Fatás, 2018).

The paper is structured as follows. Section 2 introduces the model and section 3 builds intuition for its adjustment to shocks. Section 4 describes a simulation to compute the mean output gap, while section 5 contrasts model-implied output gaps with estimates from filters. We conclude with section 6.

2 Model

Abbritti and Fahr (2013) develop a New-Keynesian DSGE model featuring search frictions in the labor market and downward nominal wage rigidity in the form of asymmetric wage adjustment costs. It is designed to capture business cycle asymmetries and largely replicates empirically documented skewness in key macroeconomic variables. As discussed in section 3, business cycle asymmetry arises endogenously as *symmetric* shocks translate into *asymmetric* model adjustments. The model is a natural choice for our exercise, since it integrates labor search and matching friction into a standard New Keynesian framework. In particular, the detailed specification of the labor market allows us to relate the model’s output gap to the output gap concept of Friedman and Phelps and to apply multivariate filters to the simulated data. This section presents a shortened summary of the model, which focuses on non-standard elements. A complete exposition can be found in the original paper.

2.1 Summary of the model

The labor market is characterized by search and matching frictions, so workers and firms need to match to create productive employment. The matching function is standard and has constant returns to scale. The model assumes contemporaneous hiring, so matches increase employment in the same period.

The representative household is a large family with members distributed along the unit interval. Each member can be employed and earn wage w_t or be unemployed and earn b_t . Consumption is pooled among family members who insure each other against the risk of becoming unemployed. The period utility function is $\log C_t$. There are risk-free nominal bonds yielding gross nominal interest rate $R_t \exp(\varepsilon_t^b)$, where $\varepsilon_t^b = \rho^b \varepsilon_{t-1}^b + \eta_t^b$ (with $\eta_t^b \sim N(0, \sigma^{2 \cdot b})$) is a zero-mean serially correlated risk premium shock. Its purpose is to introduce demand fluctuations as in Smets and Wouters (2007). Optimization yields a standard consumption Euler equation.

There are two types of firms. The first type are wholesale firms using capital and labor to produce homogeneous intermediate goods with a standard production function, in which total factor produc-

¹See, for example, Kempkes (2012), Marcellino and Musso (2011) or Kangur et al. (2019).

tivity $\exp(Z_t)$ with $Z_t = \rho^z Z_{t-1} + \eta_t^z$, $\eta_t^z \sim N(0, \sigma^{2,z})$ is subject to mean-zero productivity shocks. Because of search frictions on the labor market wholesale firms have to post vacancies to attract workers. Investment adjustment costs are symmetric, but wage adjustments give rise to asymmetric costs c_t^w described in greater detail below. The first order condition for vacancy posting is given by

$$\frac{\kappa}{\lambda_t q_t} = J_t = \alpha \varphi_t \frac{Y_t}{n_t} - w_t (1 + c_t^w) + (1 - s) \mathbb{E}_t \left[\beta_{t,t+1} \frac{\kappa}{\lambda_{t+1} q_{t+1}} \right],$$

where the LHS represents the expected costs of filling a vacancy (κ , λ_t , q_t are posting costs, the Lagrange multiplier on the household budget constraint and labor market tightness respectively). The RHS describes the expected value of a new worker, consisting of additional output $\alpha Y_t/n_t$ (expressed in units of the final good by multiplying the relative price of wholesale goods φ_t), minus wages including adjustment costs c_t^w , plus the continuation value. The first order condition for investment is standard and equates investment adjustment costs with Tobin's Q (the model abstracts from costs of capital). Wholesale firms sell their goods to retail goods firms on a competitive market.

There is a continuum of retail firms producing differentiated goods that are imperfect substitutes in the Dixit-Stiglitz consumption bundle C_t . Each firm buys wholesale goods at price $P_t \varphi_t$ and transforms them 1-to-1 and at no costs into a differentiated good variety. Monopolistic competition allows firms to charge a markup. Optimal price setting subject to standard convex adjustment costs yields the following Philipps Curve

$$\Gamma'_t \Pi_t = \epsilon (\varphi_t + \Gamma_t) - (\epsilon - 1) + \mathbb{E}_t \left[\beta_{t,t+1} \frac{Y_{t+1}}{Y_t} \Gamma'_{t+1} \Pi_{t+1} \right]$$

where Π_t is inflation, ϵ is the elasticity of substitution between retail good varieties, Γ_t adjustment costs and $\beta_{t,t+1}$ the stochastic discount factor.

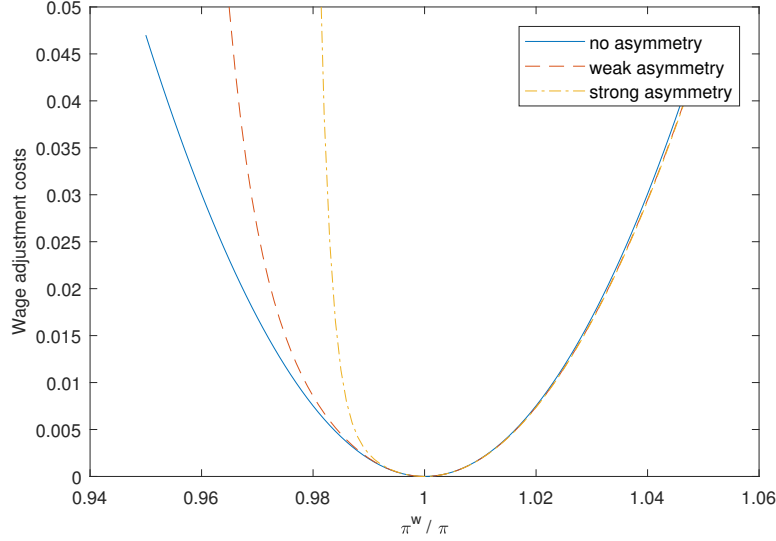
Asymmetric wage adjustment costs c_t^w give rise to downward nominal wage rigidity, the central non-standard feature of the model. It is assumed that adjustment costs c_t^w are proportional to the overall wage bill, and that they are identical for new hires and existing workers (which can be interpreted as the extension of wage agreements). Wage adjustment costs have the same functional form as in Fahr and Smets (2010):

$$c_t^w(\pi_t^w) = \frac{\phi^w - 1}{2} (\pi_t^w - \gamma)^2 + \frac{1}{\psi^2} \{ \exp[-\psi(\pi_t^w - \gamma)] + \psi(\pi_t^w - \gamma) - 1 \}$$

where π_t^w is wage inflation, which equals the the long-run deterministic output growth rate γ in the steady state.² ϕ^w governs the degree of convexity, and ψ the degree of asymmetry around steady state wage inflation. For positive values of ψ reducing wage inflation below steady state is more costly than raising it above. This asymmetry disappears as $\psi \rightarrow 0$, since the asymmetric last term then goes to zero. The graph below shows adjustment costs for three illustrative cases of no asymmetry, weak asymmetry and strong asymmetry.

²Sections 4 and 5 show detrended simulation results.

Figure 2: Illustrative wage adjustment costs



Wage determination is governed by Nash bargaining. The only non-standard element is that wage adjustment costs affect the value of a job for the firm. As shown in detail in Arseneau and Chugh (2008), the effective bargaining power of workers is given by $\omega_t = \frac{\eta}{\eta + (1-\eta)\tau_{t,t+1}}$, with η denoting the exogenous part and

$$\tau_{t,t+1} = 1 + c_t^w + \frac{\delta c_t^w}{\delta W_t} W_t + (1-s) \mathbb{E}_t \beta_{t,t+1} \left(\frac{\delta c_t^w}{\delta W_t} \right) \frac{W_{t+1}}{\Pi_{t+1}}$$

the state-dependent part.³ The effective bargaining power of workers deteriorates when wages rise ($\delta c_t^w / \delta W_t > 0$), while it increases when wages decline. Due to the asymmetry of wage adjustment costs the bargaining power rises by more in recessions (when downward wage adjustment leads to high costs) than it drops in expansions (when wages adjust upwards at comparably little costs). As stressed in Abbritti and Fahr (2013), this mechanism is an important magnifier for the impact of downward wage rigidity on the adjustment of the economy.

The resulting wage is given by

$$w_t = \omega_t \left(\alpha \varphi_t \frac{Y_t}{n_t} - c_t^w w_t + (1-s) \mathbb{E}_t [\beta_{t,t+1} J_{t+1}] \right) + (1-\omega_t) \left(b_t - (1-s) \mathbb{E}_t [\beta_{t,t+1} (1-f_{t+1}) \tilde{N}_{t+1}] \right)$$

where \tilde{N}_t is the value of employment for a family, J_t the value of a filled vacancy for a firm, and s the job destruction rate. $\omega_t c_t^w w_t$ is the deadweight loss of wage adjustment costs, which lowers the surplus of a match and thereby the wage. In this setup, the state-dependency of workers' effective bargaining power ω_t mitigates fluctuations in the wage bill. When there is upward pressures on wages, the resulting decline in bargaining power mitigates the wage adjustment. By the same token, downward pressure goes along with an increase in bargaining power, dampening the magnitude of the wage reduction.

Monetary policy is described by a standard Taylor rule with interest rate smoothing. We abstract from the monetary policy shocks that are featured in the original paper.

The resource constraint implies that output can be used for consumption, investment, and to cover adjustment costs.

³The optimal sharing rule is given by $\omega_t J_t = (1-\omega_t) \tilde{N}_t$, with \tilde{N}_t and J_t denoting the value of employment for worker and the firm respectively. It nests the standard rule when there are no wage adjustment costs, i.e. for $c_t^w = 0$ and therefore $\tau_{t,t+1} = 1$.

2.2 Output gap definition

We adopt the traditional New-Keynesian notion of the output gap (see for example Clarida, Gali and Gertler, 1999), in which potential is defined as output under price and wage flexibility, denoted by Y_t^{flex} . Using $Y_t^{rigidity}$ for output in the full-fledged model with price and wage rigidity, the gap is given by

$$gap_t = Y_t^{rigidity} - Y_t^{flex}.$$

As we will discuss further in Section 3.3, this definition corresponds to the Friedman-Phelps conception of the output gap. When the gap is closed ($gap_t = 0$), inflation is stable and the unemployment rate equals the NAIRU (taken to be unemployment under nominal flexibility). It is only when a positive (negative) gap opens up that inflation accelerates (slows down) and unemployment falls below (rises above) the NAIRU.⁴

2.3 Calibration

The only parameters we adjust relative to the baseline calibration of Abbritti and Fahr (2013) are those pertaining to exogenous shocks. The original shock calibration implies a standard deviation (SD) of gap_t of only 0.55 percent of potential output. This is unreasonably small when compared to the amplitudes of estimated output gaps, which for advanced economies are typically 3 to 4 times larger (see below). To generate more realistic output gap amplitudes in the model, we calibrate the two shock processes by matching the SD of IMF output gap estimates as well as the the SD of IMF-estimated potential around its long-run trend.⁵ The idea is to trust a leading policy institution in its assessment of the average amplitude of output gaps, which we believe is reasonable because these estimates embed the judgment of practitioners with a detailed understanding of the relevant countries. But, crucially, we *only* impose the SD of the output gap, while leaving its mean to be endogenously determined by the structure of the model.

Specifically, we use the time series of potential output and GDP from the IMF’s 2017 World Economic Outlook, which ranges from 1981 to 2017. To focus on advanced economies we use data for the G7 countries. The volatility parameter and AR(1) coefficient of demand shocks ($\sigma^{2,b}$ and ρ^b) are chosen such that the SD and first-order autocorrelation (AC) of gap_t corresponds to the moments in the WEO data when we take a non-weighted average of the G7 countries (which yields a SD of 1.77% of potential output and an AC of 0.61). Note that demand shock parameters $\sigma^{2,b}$ and ρ^b have a clear mapping into the SD and AC of the model’s output gap, because the gap is almost exclusively driven by demand shocks (es explained in the following section).

For the AR(1) coefficient of productivity shocks ρ^z we use the standard value 0.95, which makes shocks highly persistent. The parameter $\sigma^{2,z}$ governs the standard deviation of potential output around its long-run trend. To obtain a suitable empirical counterpart, we detrend IMF potential output series and compute the standard deviation of the detrended series. In the detrending procedure we first use the multiple breakpoint test embedded in eViews (using the global information criteria) to detect up to 5 breaks in a regression of potential on a time trend and a constant. Then, we allow for changes in trend growth at those break points. The standard deviation of the resulting residuals is 0.3% percent of potential (averaged over the G7 countries).

Overall, our calibration procedure implies that productivity shocks induce comparatively weak but highly persistent movements of output around its long-run trend, while demand shocks give rise to stronger but considerably less persistent fluctuations. The magnitude of output fluctuations in the model is well in line with empirical data.⁶

⁴As an aside, one could alternatively define the output gap by interpreting potential as the level of output that would occur in the model if there were no other shocks than productivity shocks (see e.g. Vetlov et al., 2011). Potential is then determined by the history of (highly persistent) productivity shocks in the model. While we will not focus on this definition, Section 3.3 shows that it is virtually identical to gap_t . In this sense our results are robust to an alternative definition of the gap.

⁵The choice of IMF estimates is arbitrary, as the dynamics of output gap estimates are very similar across various policy-making institutions. For example, the average SD of output gap estimates in G7 countries (over a comparable time span starting at the earliest available date) is 1.77% for the IMF and 1.98% for the OECD.

⁶Averaging over the G7 countries, the standard deviation of annual real GDP around its HP-filtered trend (using

3 The model’s adjustment to shocks

This section discusses the model’s adjustment to demand and supply shocks and establishes that output gaps are almost exclusively caused by demand shocks. Contractionary demand shocks cause larger gaps than an expansionary shocks of the same size. The model thus translates *symmetric* shocks into *asymmetric* output gaps. The last subsection relates the output gap in the model to the Friedman-Phelps conception.

3.1 Demand shocks

The dashed lines in Figure 3 depict the impact of a negative (i.e. expansionary) risk premium shock of one standard deviation (33 basis points in the first period). The shock reduces the return on assets held by households below the interest rate set by the central bank and thereby increases consumption. At the same time it raises investment by lowering the cost of capital. The surge in demand is met by an increase in output caused by a rise in the number of vacancies and therefore employment. The upward adjustment of wages is not constrained by the asymmetric component of adjustment costs and exceeds inflation, implying a mild increase in real wages. The central bank reacts by raising the nominal interest rate to partly offset the shock’s impact on households’ return on assets and firms’ cost of capital.

Solid lines depict the adjustment to a positive (i.e. contractionary) shock of the same size. The induced decline in output is larger and more persistent than its increase under the expansionary shock. Accordingly, the scaling-back of vacancy postings is stronger and more persistent than the hike in vacancy postings following the expansionary shock. This asymmetry is accounted for by the impact of downward nominal wage rigidity on the real wage adjustment. When the economy expands, nominal wages can rise by enough to elevate the real wage, which dampens firms’ incentives to post vacancies and therefore *mitigates* the upward adjustment of output. When the economy contracts there is a crucial difference. In this case also the real wage increases, which now reinforces firms’ willingness to reduce vacancy postings, and therefore *amplifies* the reduction in output. The real wage increases because downward nominal wage rigidity prevents nominal wages from falling by more than prices.

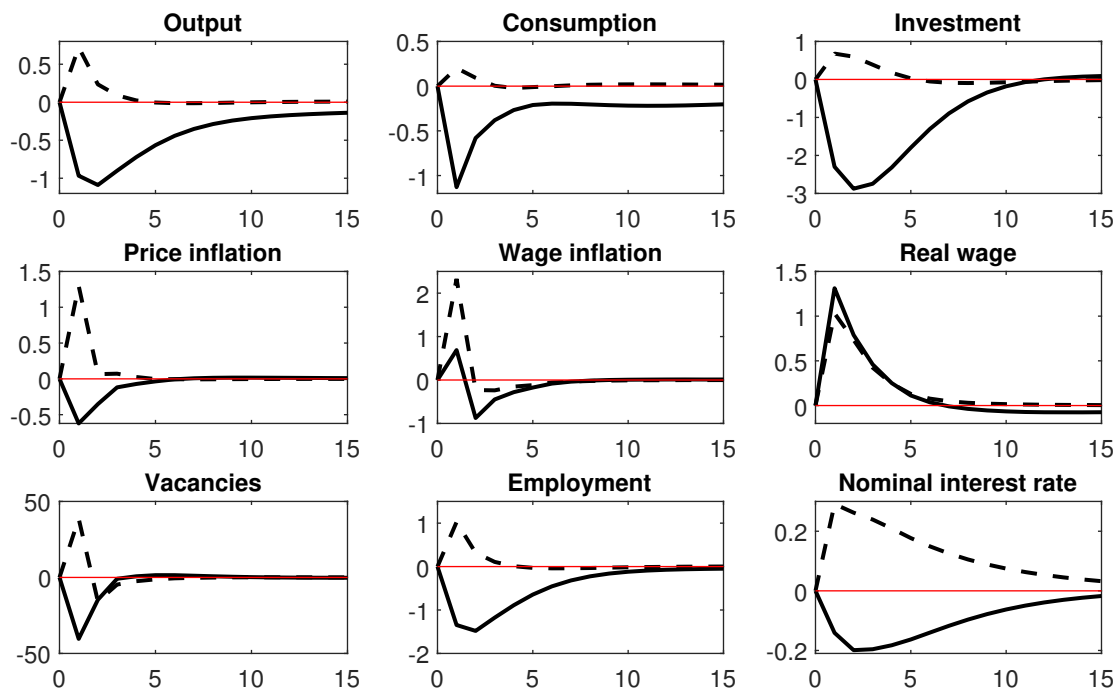
In both cases the output gap perfectly coincides with the adjustment of output shown in the first panel. Demand shocks do not affect real variables if prices and wages are perfectly flexible,⁷ and therefore leave the level of potential (i.e. output under nominal flexibility) unchanged and equal to steady state output.⁸ Percentage deviations of output from its steady state level (shown in the first panel) are thus the same as percentage deviations from potential, i.e. the output gap. Comparing the gap across both scenarios, we observe that the negative gap caused by a contractionary shock is larger in magnitude and more persistent than the positive gap caused by an expansionary shock of the same size.

$\lambda = 6.25$ and a sample from 1980-2019) is 1.23 percent of the trend. Applying the same calculus to annualized GDP in the model yields a standard deviation of 1.02 percent.

⁷Under nominal flexibility output cannot deviate from its steady state value, because any deviation would cause an indefinitely strong price adjustment that in turn leads to a central bank reaction counteracting the deviation of output. An expansionary demand shock, for example, causes an instantaneous upward-adjustment in prices that leads to an increase in the nominal rate which perfectly offsets the shock.

⁸The steady state allocation under nominal rigidity equals the flexible-price/flexible-wage allocation, because all nominal variables had infinite time to adjust. Actual output thus equals potential output and the gap is closed.

Figure 3: Adjustment to positive (dashed lines) and negative (solid lines) risk premium shock of one standard deviation.

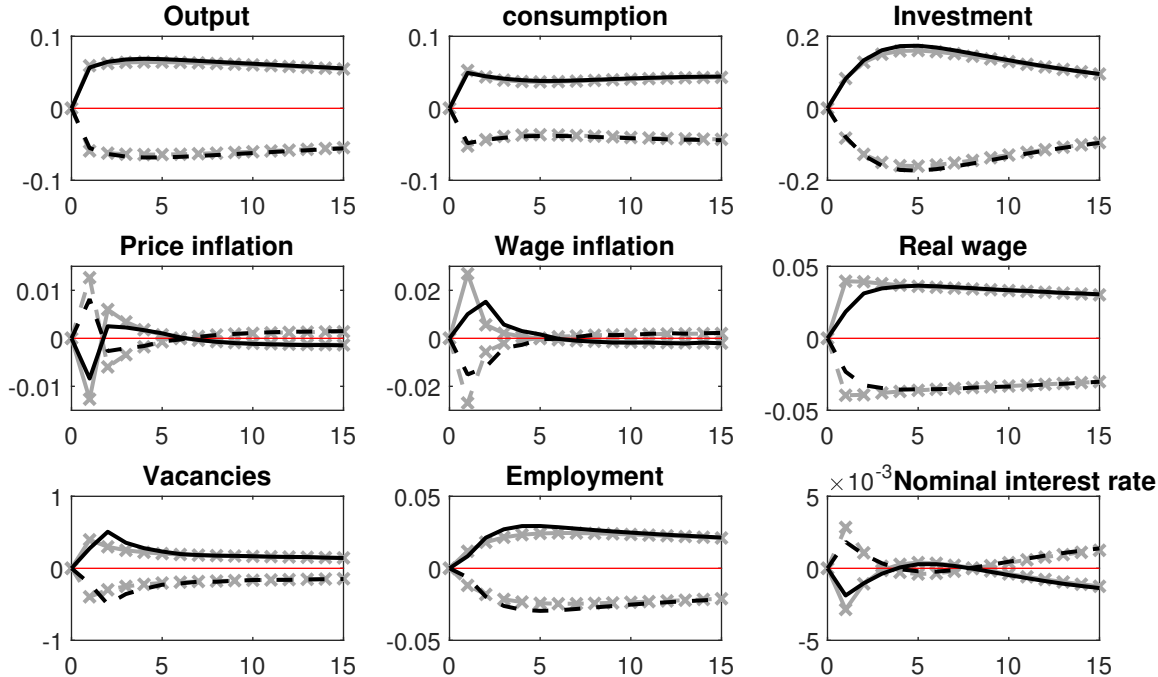


3.2 Supply shocks

The solid (dashed) black lines in Figure 4 depict the model's adjustment to a positive (negative) productivity shock of one standard deviation. Changes in productivity give rise to moderately large low-frequency movements of output rather than inducing strong volatility at high frequencies. Compared to demand shocks, the impact of productivity shocks on output is therefore small but highly persistent. A positive shock raises productivity and increases the value of a worker for the firm, which leads firms to post more vacancies and increase employment. Investment surges because of the higher marginal product of capital, and consumption rises due to a positive wealth effect. The increase in the real wage – resulting from wage inflation combined with a mild decline in prices – offsets some of the uptick in vacancy postings caused by the productivity gain. The adjustment of the nominal interest rate is negligible due to the small size of price adjustments. We skip the discussion of the negative productivity shock because the model does not exhibit significant asymmetries in the adjustment to shocks of opposite signs. Asymmetry is missing because the adjustment of consumption and investment is dominated by the direct impact of the shock. Nominal rigidity, which is the only asymmetric feature of the model, thus plays a negligible role.

Grey lines marked with crosses depict the model's adjustment when prices and wages are flexible. The two grey lines in the first panel thus represent the adjustment of potential output to a positive and negative productivity shock. The shocks have a persistent impact on the level of potential, which makes intuitive sense as they affect the productive capacity of the model economy. The output gap is the distance between the black and grey lines in the first panel, and it is negligibly small for both the positive and the negative shock. Again, the explanation lies in the fact that output adjustment is dominated by the the direct impact of the shock and is therefore only to a very small extent affected by nominal rigidity (which is the only reason for potential to deviate from actual output).

Figure 4: Adjustment to positive (solid black lines) and negative (dashed black lines) productivity shock of one standard deviation. Grey lines with crosses show the adjustment under nominal flexibility (solid for a positive and dashed for a negative shock).



3.3 Relation of the model output gap to the Friedman-Phelps ‘natural rate’

We now examine how the model-implied output gap – derived from within a New Keynesian framework – relates to the ‘natural rate’ of unemployment first introduced to the literature by Friedman (1967) and Phelps (1967) and alternatively referred to as the ‘non-accelerating inflation rate of unemployment’ or NAIRU. The NAIRU in the context of the model is interpreted as the level of unemployment arising from labor search and matching frictions when prices and wages are flexible. This friction-related unemployment arises even in the absence of cyclical swings (when the output gap is closed), and it declines (increases) when productivity shocks expand (reduce) the productive capacity of the economy.⁹

Two observations are in order to organize our thinking on the relation in the model between output gap, unemployment gap (defined as the difference between unemployment and NAIRU) and inflation. First, we know from the previous subsection that the model adjustment to a productivity shock is dominated by its direct impact, with price and wage adjustment being negligible. Consequently, nominal rigidity has no significant role to play, and no gaps open up as output and unemployment move 1-to-1 with potential and the NAIRU respectively (the grey and black lines in the output and employment panels of Figure 4 effectively overlap). Therefore we can focus exclusively on the role of demand shocks in affecting the output gap. Second, as indicated in Subsection 3.1, demand shocks do not affect potential output and the NAIRU. The deviation in the output panel of Figure 3 thus shows the output gap, while a positive (negative) deviation in the employment panel indicates that unemployment falls below (rises above) the constant NAIRU.

Figure 3 shows that inflation accelerates (decelerates) when the output gap turns positive (negative). We also observe that unemployment rises (falls) below the NAIRU if and only if there is a negative (positive) output gap. This implies that there is a unique level of unemployment – corresponding

⁹This can be seen in the grey lines in the employment panel of Figure 4. Positive productivity shocks increase employment under nominal flexibility while negative shocks reduce it. Since the labor force is constant, unemployment moves in the opposite direction.

to the flexible price/wage allocation and to a closed output gap – at which inflation is stable. At any other rate of unemployment, there is either a positive or a negative output gap, and inflation accordingly accelerates or declines. Thus the flexible price/wage unemployment level of the New Keynesian framework with labor search frictions may be identified with the NAIRU, or natural rate. Any deviations from the natural rate imply that inflation is either rising or falling.¹⁰

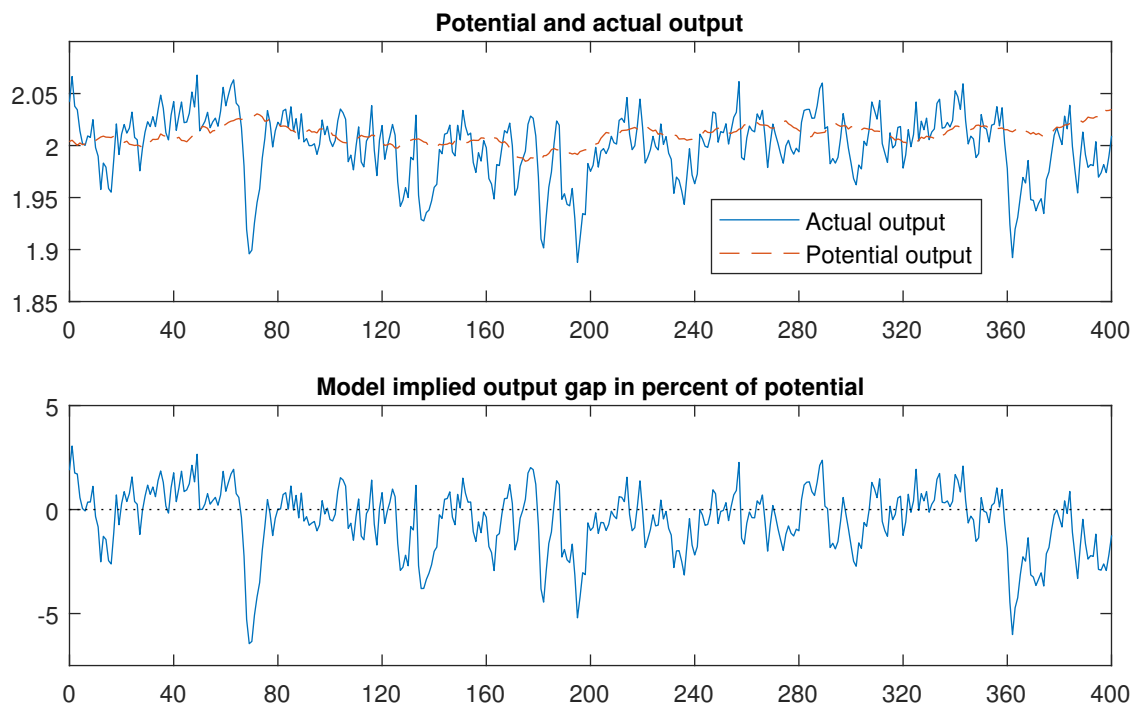
4 Mean output gap in the model

This section outlines the use of the model to simulate data and discusses the properties of the output gap in this synthetic data. The first step in the simulation is to draw time series for demand and supply disturbances from the respective distributions described in section 2. Feeding these shock series into the model, we generate time series for all endogenous variables. We do so twice, using the same shock series but different calibrations of the model: First, we use the baseline calibration (with nominal price and wage rigidity) and call the obtained output series ‘actual output’. Second, we deactivate nominal rigidity by setting price and wage adjustment costs to zero. The resulting output series is referred to as ‘potential output’, in line with the New-Keynesian definition of the output gap. We generate 400,000 quarters of data and, following Abbritti and Fahr (2013), use a second order accurate perturbation method to account for the model’s asymmetries.

For illustrative purposes Figure 5 shows an arbitrarily chosen sub-sample of 400 quarters of simulated output data. We abstract from the deterministic long-run productivity trend in the model, so actual and potential output (shown in the upper panel) fluctuate around the steady state output level of around 2. Changes in potential output are driven by productivity shocks and are therefore highly persistent and relatively small. Actual output, by contrast, is also affected by relatively large and short-lived demand shocks, and thus exhibits stronger and more frequent fluctuations. We observe that deviations of actual output above potential tend to be smaller than those below potential; as explained above, negative demand shocks impact output more strongly than positive ones. Accordingly, the output gap shown in the lower panel (the distance between actual and potential output expressed as percentage of potential) appears to be negative more often than positive, and the negative deviations tend to be larger in magnitude. Over the full sample of 400,000 periods, the gap averages -1.02 percent of potential output. Again, note that this negative gap is driven entirely by the asymmetric wage adjustment costs, as the underlying shock series are symmetric.

¹⁰Intuitively, under nominal flexibility firms and workers are unconstrained and set prices and wages such that they have no incentive to further adjust them. The resulting level of unemployment is thus associated with the absence of price and wage adjustment pressures.

Figure 5: Simulated time series for actual output and (model-implied) potential.



The full-sample mean of the output gap masks strong variation across sub-samples, particularly between ‘calm’ periods where deviations of actual output from the steady state are relatively small, and ‘turbulent’ periods where these deviation are relatively large. In the following exercise we study this variation by splitting the sample into turbulent and calm periods. To this end a 3-year moving average of the deviation of output from its steady state level is moved along the sample. Whenever the absolute value of this moving average exceeds a certain threshold, the respective time window counts towards the turbulent sub-sample. Conversely, when the average is below that threshold, the window is counted towards to the calm sub-sample. The threshold is chosen such that the summed duration of turbulent periods accounts for a certain fraction of the total sample. For example, the grey areas in Figure 6 (showing an illustrative sub-sample) indicate turbulent periods, where the threshold is chosen such that one third of the full sample is counted as turbulent. In this case the output gap averages -1.9% of potential in the turbulent third of the sample, and -0.6% percent in the remaining, calm two-thirds.

Figure 6: Selection of sub-samples

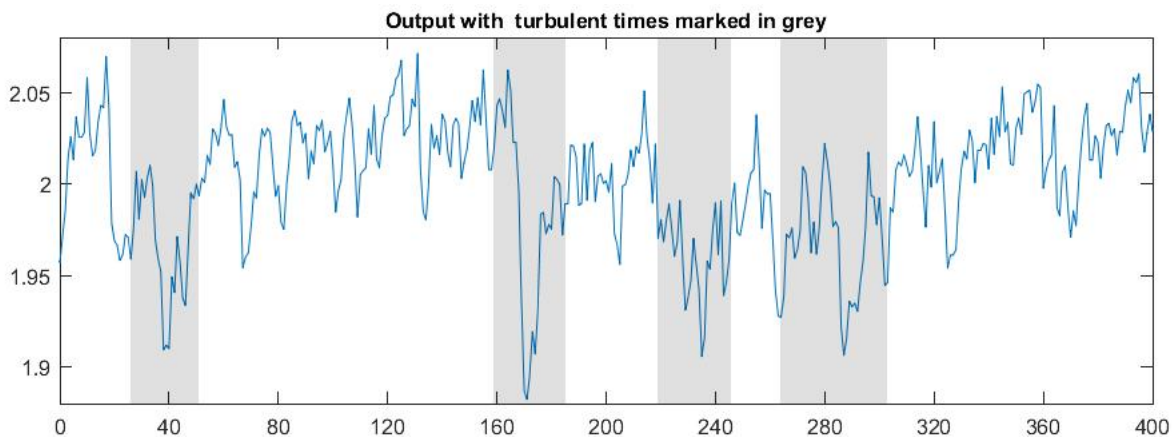


Table 1 presents results for different splits of the sample. For example, if one quarter of the sample is selected as turbulent (2nd row), the mean gap in this sub-sample amounts to -2.68% of potential, while the mean gap in the remaining part of the sample is -0.46%. The magnitude of the negative output gap is larger the smaller the turbulent sub-sample, ranging from -3.61% when only 10% of the sample is selected to -1.95% when we select 50%. The explanation is straightforward, and follows from the functional form of the wage adjustment costs. As Figure 2 makes clear, the asymmetry between positive and negative wage adjustment costs is greater the larger the shock. There is little difference in wage adjustment costs between a small positive shock and a small negative shock, but the difference for larger shocks is much greater. A large negative demand shock is therefore amplified by the disproportionately greater wage adjustment cost, which results in larger downward adjustments to employment and output. The economic intuition for the functional form of the adjustment costs is that a small negative shock does not necessarily require a cut to nominal wages; instead nominal wages can simply increase by less than price inflation. This is easier (i.e. less costly) for an employer than nominal wage cuts (as evidenced by the spike at zero in Figure 1). But as the size of the negative shock increases, nominal wages need to fall and this is much costlier for employers than a nominal wage increase of equal magnitude.

Thus, as a result of asymmetric wage adjustment costs, large deviations of output from steady state tend to be negative and accompanied by negative output gaps. The tighter the selection criterion for turbulent times (i.e. the higher the selection threshold), the more narrowly are turbulent sub-periods centered around the largest periods of negative gaps – resulting in a more negative average.

Focusing on the mean gap in calm periods shows that the full-sample mean of -1.03% is predominantly driven by relatively rare periods of large negative gaps. The average gap excluding the more turbulent half of the sample is close to zero (-0.08%), and excluding the most turbulent third and quarter of the sample leads to mean gaps of -0.33% and -0.46% respectively. This is reassuring in the sense that the conventional notion of a zero-mean output gap is approximately true at least about half the time. However, as we will discuss in the following section, the large negative gap during turbulent times severely biases conventional output gap estimates during those periods when an accurate estimation of the output gap is likely most important for policy.

Table 1: Average output gap in sub-samples

Sample split	Mean output gap
Turbulent/Calm	in sub-samples
10% / 90%	-3.61 / -0.72
25% / 75%	-2.68 / -0.46
33% / 67%	-2.40 / -0.33
50% / 50%	-1.95 / -0.08
Full sample	-1.03

Output gaps are reported in percent of potential output.

5 Accuracy of output gap filters

We have argued that downward nominal wage rigidity leads to a negative mean output gap, and have demonstrated that it does so in the model. At the same time, standard output gap estimation methods (referred to as ‘filters’) have the property that the estimated gap has a long-run mean of zero. In this section we show that this discrepancy can lead to a bias in output gap estimates. Examining the accuracy of filters is straightforward when we exploit the fact that the output gap in the model is observable. Using the time series of actual and potential output studied in the previous section allows a straightforward assessment of the accuracy of univariate filters that use only output: the filter can be applied to actual output and the resulting estimate of potential can be compared to the true level of potential obtained from the model under nominal flexibility. For multivariate filters we also generate time series for additional model variables.

Studies of the reliability of output gap estimates typically focus on real-time estimates and benchmark them against final estimates, which are interpreted as the best available estimate of the true output gap (for a seminal contribution see Orphanides and Norden, 2002).¹¹ Our strategy is different in that our benchmark is not the final output gap estimate, but the true output gap in the model, which is observable in the simulated data. While our approach relies on a simplified model of the economy, it alleviates the principal flaw of existing studies, which is that they lack a benchmark of the true output gap.

We focus on two illustrative filtering methods, the Hodrick-Prescott (HP) filter and the IMF’s multivariate filter. The first is the most widely used among the category of purely statistical filters, while the latter will serve as an example of filters that incorporate structural economic relationships into the estimation. Our findings, however, apply in principle more broadly to all zero-mean filters, including e.g. the band-pass filter or production function based methods that use zero-mean filters to estimate trend components for input factors and TFP.¹² Subsections 5.1 and 5.2 discuss the zero-mean property of both filters and examine in detail their biases, which are summarized in Table 2 (we refer to this table throughout the section). A corollary of the bias is that filters mis-estimate the point in time when the sign of the output gap switches. Subsection 5.3 provides some intuition for why unwarranted symmetry in the filters – manifested in a zero mean estimate – leads to a spurious interpretation of negative demand shocks as affecting potential, which is at the core of the bias.

¹¹Real-time estimates are based on data available at the time of the estimation, while final estimates use the most recent data vintage.

¹²Band-pass filtering of historical real GDP data for G7 countries (with samples from the earliest available data on FRED to 2018Q2, and using 6 to 32 quarters as business cycle frequency) yields an average output gap of -0.05%. The European Commission’s production function based estimates for all countries in the AMECO database average to -0.16% (from the earliest available data to 2019).

Table 2: Average output gap in sub-samples: model-implied versus estimated

Sample split Turbulent/Calm	True	HP Filter		Multivariate Filter**	
	Avg. Gap	Avg. Gap	Bias*	Avg. Gap	Bias*
10% / 90%	-3.61 / -0.72	-0.34 / 0.03	3.26 / 0.76	-1.83 / 0.22	2.05 / 0.91
25% / 75%	-2.68 / -0.46	-0.15 / 0.05	2.52 / 0.51	-1.16 / 0.40	1.71 / 0.79
33% / 67%	-2.40 / -0.33	-0.13 / 0.06	2.27 / 0.39	-0.91 / 0.46	1.61 / 0.73
50% / 50%	-1.95 / -0.08	-0.09 / 0.09	1.85 / 0.17	-0.56 / 0.58	1.43 / 0.61
Full sample	-1.03	0.00	1.03	0.00	1.03

All values are in percent of potential output.

** Biases are measured as estimated minus true average gap, so positive values indicate that the estimated gap is on average more positive than the true gap.*

*** Output gap estimates for the MVF are derived using annualized data.*

5.1 The Hodrick-Prescott Filter

The HP filter is often used to decompose an output series into its trend component (interpreted as potential) and its cyclical component (interpreted as output gap). We prove in the appendix that the average of the cyclical component is always zero, which we refer to as the zero mean property of the filter. For the HP filter, the zero mean property holds not just in the limit, but for any sample size. The upper panel in Figure 7 again shows an illustrative sub-sample with simulated actual output (blue), true potential in the model (orange) as well as HP-estimated potential (yellow), i.e. the trend component from an HP filtering of simulated actual output data using $\lambda = 1600$ for quarterly frequency (as suggested by Ravn and Uhlig, 2002).

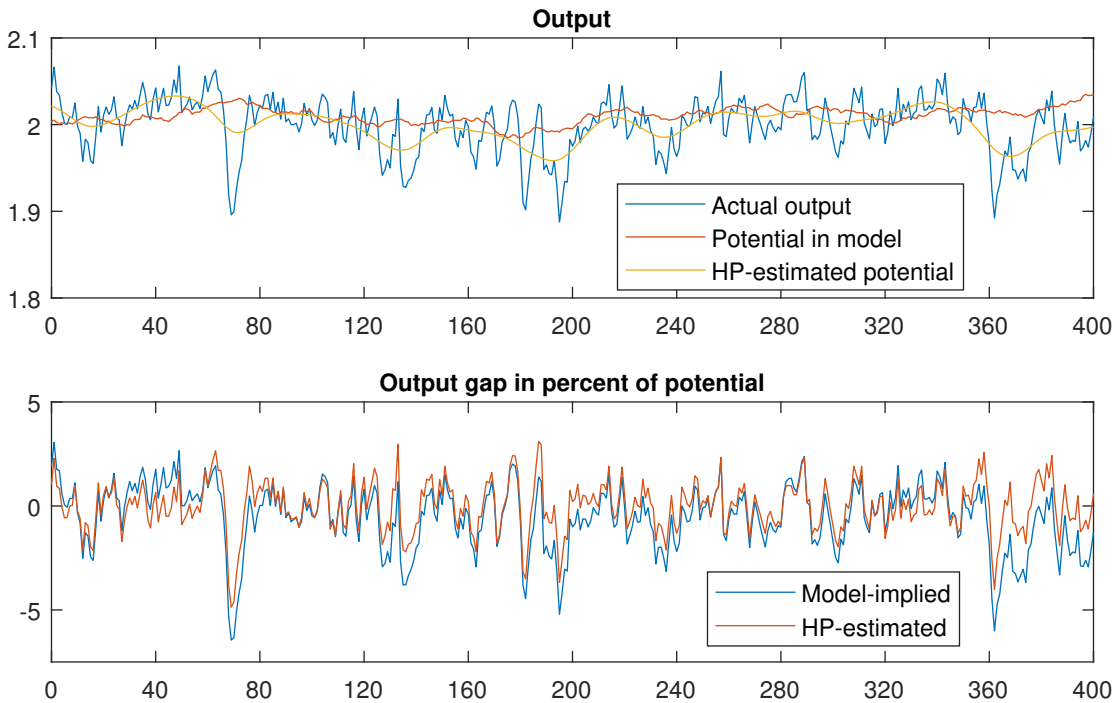
Estimated potential tends to lie below true potential when the true output gap is negative, i.e. when actual output is below true potential. The reverse holds in periods of a sustained positive gap – the estimate then overshoots true potential, as e.g. between periods 20 and 60 – but positive gaps are less frequent and tend to be smaller than negative gaps. This asymmetry leads to a quantitatively significant bias: while the HP-estimated gap has a mean of zero over the whole sample, the true gap averages -1.02 percent of potential. The difference, i.e. -1.02 percent, constitutes the filter’s average bias. To put this another way, the construction of the HP filter implies that estimated potential must ‘cut through’ actual output such that cumulative deviations are zero. It is thus centered around actual output, while true potential is larger.

Table 2 reports the average bias when the sample is split into turbulent and calm sub-samples as in the previous section. The bias is smaller in calm than in turbulent times. If we exclude the most turbulent third of the sample (third row), the true output gap averages -0.33 percent while the HP estimates have a mean of 0.09 percent. The resulting bias is 0.39 percent on average, which is less than half as large as for the full sample. By contrast, if we focus on the most turbulent quarter of the sample, the bias amounts to 2.52 percent on average. The larger bias in turbulent times is explained by the behaviour of estimated potential shown in Figure 7. It is during large negative deviations of output from potential when the bias is largest, and the turbulent sub-sample encompasses precisely those periods. The bias is thus predominantly driven by periods of strong macroeconomic turbulence, in particular severe recessions. This is reassuring in the sense that the bias is relatively weak most of the time. However, it is large during deep recessions, suggesting that over-reliance on statistical filters during such episodes could create a danger of policy under-reaction.

The tendency of the HP filter to under-estimate the true level of potential also affects the timing of when the output gap changes sign, i.e. when the time series for estimated potential intersects with actual output. Consider an upswing, when both the true and the estimated gap turn from negative to positive. The estimated gap tends to be above the true gap and therefore mechanically cuts through zero (from below) ahead of the true gap; the filter prematurely signals that the economy has begun to operate above potential. By the same token, during a downturn, an estimated output gap above the true gap will intersect zero (from above) too late, leading to a delay in signaling that output has fallen

below potential. To quantify this bias, we apply the Harding-Pagan business cycle dating algorithm to partition the simulated data into periods of cyclical upswings and downturns.¹³ A switching of the true output gap from negative to positive (during upswings) is on average preceded by a switching of the estimated output gap 4.7 quarters earlier. A switching of the true output gap from positive to negative (during downturns) is on average followed by a switching of the estimated gap 0.5 quarters later.¹⁴

Figure 7: Accuracy of HP filter.



5.2 The IMF's multivariate filter

The IMF's multivariate filter (MVF), described in Blagrove et al. (2015), is a prominent example of a more elaborate filtering methodology that is not purely statistical but features economic structure. The MVF embeds an Okun's law relationship as well as a Phillips curve to inform the estimation of the output gap. In contrast to the HP filter, estimates obtained from the MVF do not necessarily have a mean of zero for all possible data samples. However, MVF estimates have a mean of virtually zero (0.006%) for our simulated data, and also for historical data for the G7 countries.¹⁵ The output gap in the filter is modelled as $gap_t = \phi gap_{t-1} + \epsilon_t$ ($0 < \phi < 1$) with a prior distribution of ϕ that is typically centred around a mode of 0.6. As a result gap_t tends to gravitate strongly towards its unconditional mean of zero.¹⁶ This makes it difficult to accommodate the more persistent negative output gaps which tend to arise in the presence of downward nominal wage rigidity. So while the zero-mean property is not as absolute as in the case of the HP filter, it is pronounced enough to give rise to a significant bias.

¹³The calibration of the algorithm is chosen such that the average duration of a full cycle is about 8 years.

¹⁴When the gaps switch signs more than once during a given upswing or downturn, we consider only the first sign switching. The dating algorithm can identify cycles during a sustained negative deviation of output, which produces upswing periods that do not see the true gap turning positive. These episodes are excluded from the analysis.

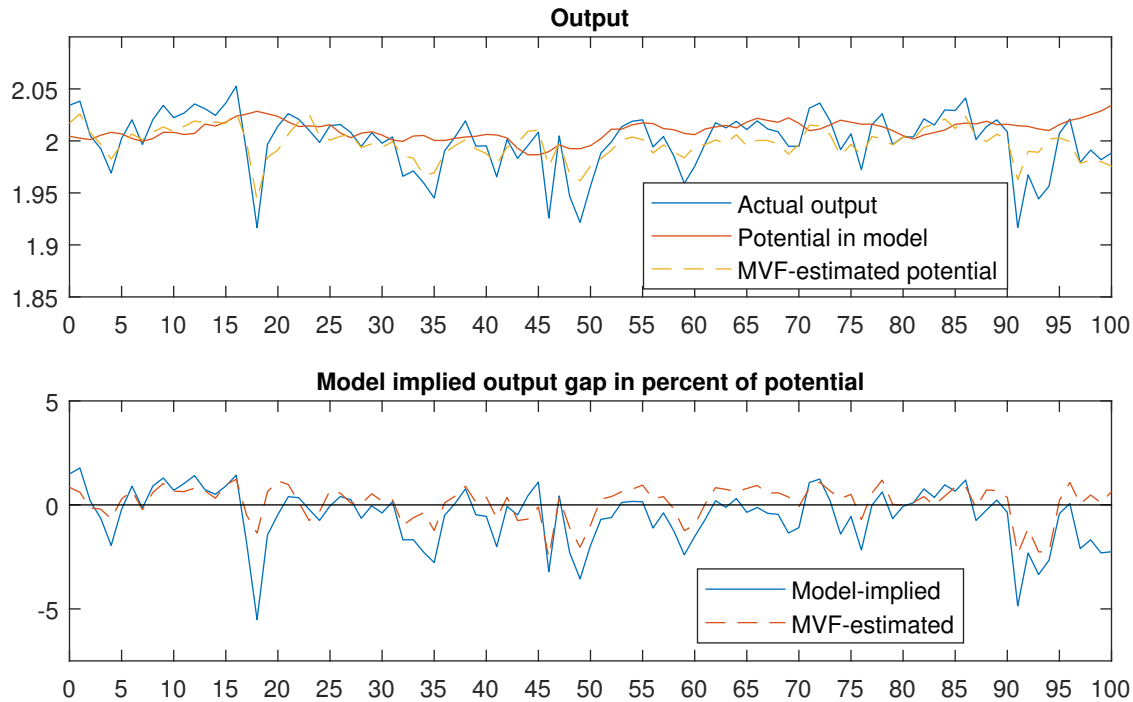
¹⁵We applied the MVF to historical data for the G7 countries, with samples starting at the earliest available observation and going to 2018. The average output gap estimate is -0.08%.

¹⁶Taking expectations yields $\mathbb{E}[gap] = \phi \mathbb{E}[gap] + \mathbb{E}[\epsilon]$, which implies $\mathbb{E}[gap] = 0$ because $\mathbb{E}[\epsilon] = 0$.

Since the filter is built for an annual frequency, we annualized the simulated quarterly time series for the three variables required to run the filter: output, the unemployment rate, and inflation.¹⁷ Figure 8 again shows the illustrative sub-sample, this time including the MVF-estimated potential output and the output gap. As in the case of the HP filter, MVF-estimated potential tends to be below true potential, especially when there are large and persistent negative output gaps. The last two columns of Table 2 report average gap estimates and the corresponding bias when the sample is split into turbulent and calm sub-samples as before. While the average bias over the full sample is the same as for the HP filter, the MVF performs significantly better in turbulent periods. For example, the bias in the most turbulent tenth of the sample amounts to 2.05 percent, whereas it is 3.26 percent for the HP filter. Incorporating economic structure appears to help in identifying large negative gaps, but the overall finding – a significant bias in times of economic distress – still holds. In calm times the filter spuriously estimates a modest positive output gap, e.g. 0.58 percent on average during the calm half of the sample.

The bias in the estimated level of potential output translates into errors in the estimated point in time when the output gap closes, as explained in the previous sub-section. We again apply the Harding-Pagan business cycle dating algorithm to separate the sample into upswings and downturns. During upswings the estimated gap switches from negative to positive 1.2 years too early (relative to the true output gap), while in downturns it switches from positive to negative 0.7 years too late.¹⁸

Figure 8: Accuracy of IMF’s multivariate filter.



5.3 The zero-mean property of filters and shock identification

In the presence of downward nominal wage rigidity, negative deviations of output from potential sum to a larger integral than positive deviations from potential. For a filter that is constrained by the zero-mean property, the positive areas in the integral therefore do not create enough ‘space’ to accommodate the more substantial negative deviations of output from true potential, with the

¹⁷Every four consecutive quarters of simulated data are averaged into one year of annual data.

¹⁸The calibration of the dating algorithm (for the annualized data) is again chosen such that the average duration of a full cycle is about 8 years.

result that potential output cannot be correctly estimated. The only way for zero-mean filters to accommodate the asymmetry caused by downward nominal wage rigidity is to 'shift down' estimated potential below its true level (thereby making negative areas in the integral smaller and positive areas larger) such that the integral becomes zero. In other words, the asymmetry is picked up as a spurious reduction in estimated potential.

This downward bias in estimated potential is manifested in a tendency of zero-mean filters to incorrectly interpret negative demand shocks as supply shocks (reducing potential) and positive supply shocks as demand shocks (not lifting potential). There is a true sequence of supply and demand shocks that is compatible with the true series of potential output. Since the zero-mean property forces filters to estimate a lower level of potential, the estimated series of shocks has to exaggerate negative supply shocks and understate positive supply shocks. In our model the most prominent consequence of this flaw is that prolonged negative demand shocks are interpreted as lowering potential, to an extent that increases with the severity and duration of the shock. The intuition is straightforward: if large and prolonged negative demand shocks were correctly interpreted as such, they would open up large negative areas in the integral between actual and potential output. But this is incompatible with the zero-mean property.¹⁹

This critique serves to illuminate the recent empirical analysis by Coibion, Gorodnichenko and Ulate (2019). The authors examine potential output estimates by various institutions and report that econometrically identified demand shocks are incorporated into estimated potential to an extent that increases over time. For example, monetary shocks and shocks to government spending are reported to have transitory effects on GDP, but nevertheless lead to a gradual adjustment of estimated potential in the same direction. To the extent that the filters underlying those estimates have the zero-mean property, our findings provide an explanation for this observation.

Our analysis abstracts from hysteresis effects, which could cause potential output to decline as a result of long-lasting negative demand shocks. But it would require an extremely unlikely coincidence for hysteresis effects to exactly cancel out the asymmetry induced by downward nominal wage rigidity, leaving the output gap centered at precisely zero. Indeed, our work shows that the repeated downward revisions of estimated potential since 2008 are plausibly a consequence of the zero-mean property of filters; this should be taken into account when the revisions are instead interpreted as evidence of hysteresis.²⁰

6 Conclusion

There is no theoretical reason to insist that positive deviations from potential output must always equal negative deviations over time, that is, that the mean output gap must be zero. Indeed, Okun's 1962 paper conceived of the output gap (the 'GNP gap' as he called it) as intrinsically negative; with employment sometimes falling below full employment but not above it. The same holds in Milton Friedman's 'plucking model' that was recently formalized by Dupraz, Nakamura and Steinsson (2017). This simply enshrined the intuition that while unemployment is a commonplace, observable phenomenon, 'overemployment' is a much more theory-laden concept. We would argue that this remains the case. Presumably there are sharp physical limits to the number of extra hands that can usefully be employed with a given stock of capital, capping the magnitude of positive deviations from potential output relative to negative deviations.

¹⁹To further illustrate this point, consider an economy that exists for two periods and is only subjected to symmetric, one-period demand shocks that leave potential unaffected. Suppose there is a negative shock realization in the first period and a positive realization of the same size in the second. Due to asymmetric shock adjustment, the negative shock causes output to decline by -2% in the first period, while the positive shock causes output to increase by only 1% in the second period. The output gaps on both periods are -2% and 1% respectively, and the integral between output and potential is $-2\%+1\%=-1\%$. Zero-mean filters have to interpret the negative shock to some extent as a supply shock (lowering potential), because the only way to bring about an integral of zero is to estimate that potential output has declined. For example, a filter could incorrectly interpret half of the output decline in the first period as being caused by a supply shock. Potential in the first period would then be lower by 1% (while it is unchanged in the second period, abstracting for simplicity from persistence). This shrinks the estimated gap in the first period to -1%, which means that the integral is $-1\%+1\%=0\%$ and the zero-mean property fulfilled.

²⁰Of course this argument does not apply to micro-econometric evidence for hysteresis effects.

In this paper, however, we make no appeal to such arguments. Instead we demonstrate that downward nominal wage rigidity in a standard, otherwise symmetrical New Keynesian model with labor search frictions, inevitably leads to greater adjustment of output and employment in downturns than in booms. If positive and negative demand shocks are distributed symmetrically, this implies a negative mean output gap. Moreover, this output gap is precisely the object of relevance to policymakers – in the Friedman-Phelps sense – despite its New Keynesian definition in terms of distance from the flexible price allocation.

We further show that common filtering techniques are intrinsically biased in a world characterized by asymmetric shock adjustment and a negative mean output gap. Both purely statistical filters such as the HP filter, and other techniques that incorporate greater economic structure, such as the IMF's MVF, tend to impose a zero mean on the cyclical component which is extracted from the data. This zero mean property causes the filters to erroneously underestimate potential output, especially during severe downturns.

Our work suggests that the best way forward for estimating cyclical deviations from potential output lies in methods that do not – implicitly or explicitly – incorporate a zero-mean property. For example, Coibion, Gorodnichenko and Ulate (2019) suggest stacking together identified supply shocks (using the identification strategy of Blanchard and Quah, 1989) to construct a series for potential that is not stationary around actual output. This approach is not subject to our critique, but requires choosing a starting point in time when the output gap is closed.

Our argument also indicates an urgent need for research on optimal policy in a negative mean output gap world, as opposed to the zero mean output gap setting implicitly or explicitly assumed by policymaking institutions. As one example, if an economy spends more time below full employment than above it, then fiscal policy aiming to be symmetrical around full employment will be expansionary more often than contractionary. Thus, to avoid explosive debt dynamics, fiscal consolidation may need to begin even while the output gap remains negative. As another example, Taylor rules for monetary policy are likely to differ significantly if time-series for the output gap are allowed to have negative means rather than having a zero mean imposed through filtering.

More fundamentally, our work here illustrates the perils for macroeconomic policy of relying too heavily on output gap estimates. If the economy is characterized by a negative mean output gap, as we argue, then standard filtering technique will be intrinsically biased, and the bias will be greater during the turbulent periods when stabilization policies are most required. This at least suggests that policy should place greater reliance on observables, such as inflation, unemployment and the level of public debt.

Appendix: Proof that the cyclical component of the HP filter has mean zero

We first show that the HP filter can be written in a linear form. The filter computes the trend component $\{\hat{m}_t\}$ of a time series $\{y_t\}$ as solution to the following minimization problem:

$$\arg \min_{\{m_t\}} \sum_{t=1}^T (y_t - m_t)^2 + \lambda \sum_{t=3}^T (m_t - 2m_{t-1} + m_{t-2})^2$$

which, in vector notation, can be written as

$$\arg \min_{\mathbf{m}} (\mathbf{y} - \mathbf{m})' (\mathbf{y} - \mathbf{m}) + \lambda (\mathbf{B}\mathbf{m})' (\mathbf{B}\mathbf{m})$$

$$\text{with } \underset{(T \times 1)}{\mathbf{y}} = \begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix}, \underset{(T \times 1)}{\mathbf{m}} = \begin{pmatrix} m_1 \\ \vdots \\ y_T \end{pmatrix} \text{ and } \underset{(T-2 \times T)}{\mathbf{B}} = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}.$$

Setting the first derivative to zero yields the following solution for the vector of the estimated trend, $\hat{\mathbf{m}}$:

$$\hat{\mathbf{m}} = (\mathbf{I} + \lambda \mathbf{B}' \mathbf{B})^{-1} \mathbf{y} \quad (1)$$

Note that $(\mathbf{I} + \lambda \mathbf{B}' \mathbf{B})^{-1}$ is symmetric because \mathbf{B} is.

In the following we prove that the estimated cyclical components $\hat{u}_t = y_t - \hat{m}_t$ have a mean of zero. Let $\mathbf{1} = (1 \cdots 1)'$ denote the unit vector and note that $\mathbf{B}\mathbf{1} = \mathbf{0}$ because all rows in \mathbf{B} sum up to zero. It also holds that $(\mathbf{I} + \lambda \mathbf{B}' \mathbf{B}) \mathbf{1} = \mathbf{1}$, which can be verified by rewriting the LHS as $\mathbf{I}\mathbf{1} + \lambda \mathbf{B}' \mathbf{B}\mathbf{1} = \mathbf{1} + \lambda \mathbf{B}' \mathbf{0} = \mathbf{1}$. Multiplying from the left by $(\mathbf{I} + \lambda \mathbf{B}' \mathbf{B})^{-1}$ and transposing yields

$$\mathbf{1}' = \mathbf{1}' (\mathbf{I} + \lambda \mathbf{B}' \mathbf{B})^{-1}. \quad (2)$$

The mean of estimated trend components is given by $\bar{\hat{m}}_t = \frac{1}{T} \sum_{t=1}^T \hat{m}_t = \frac{1}{T} \mathbf{1}' \hat{\mathbf{m}}$. Using (1) it can be written as $\bar{\hat{m}}_t = \frac{1}{T} \mathbf{1}' (\mathbf{I} + \lambda \mathbf{B}' \mathbf{B})^{-1} \mathbf{y}$, which simplifies to $\bar{\hat{m}}_t = \frac{1}{T} \mathbf{1}' \mathbf{y} = \bar{y}_t$ using (2). It follows that the mean of the estimated cyclical component is zero: $\bar{\hat{u}} = \frac{1}{T} \sum_{t=1}^T \hat{u}_t = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{m}_t) = \frac{1}{T} \sum_{t=1}^T y_t - \frac{1}{T} \sum_{t=1}^T \hat{m}_t = \bar{y}_t - \bar{\hat{y}}_t = 0$.

References

- Abbritti, Mirko, and Stephan Fahr.** 2013. “Downward wage rigidity and business cycle asymmetries.” *Journal of Monetary Economics*, 60(7): 871–886.
- Arseneau, David M, and Sanjay K Chugh.** 2008. “Optimal fiscal and monetary policy with costly wage bargaining.” *Journal of Monetary Economics*, 55(8): 1401–1414.
- Ball, Laurence.** 2014. “Long-term damage from the Great Recession in OECD countries.” *European Journal of Economics and Economic Policies: Intervention*, 11(2): 149–160.
- Benigno, Pierpaolo, and Luca Antonio Ricci.** 2011. “The inflation-output trade-off with downward wage rigidities.” *American Economic Review*, 101(4): 1436–66.
- Blagrove, Patrick, Mr Roberto Garcia-Saltos, Mr Douglas Laxton, and Fan Zhang.** 2015. “A simple multivariate filter for estimating potential output.” , (WP/15/79).
- Blanchard, Olivier, and Danny Quah.** 1989. “The Dynamic Effects of Aggregate Demand and Supply Disturbances.” *The American Economic Review*, 79(4): 655–673.
- Clarida, Richard, Jordi Gali, and Mark Gertler.** 1999. “The science of monetary policy: a new Keynesian perspective.” *Journal of economic literature*, 37(4): 1661–1707.
- Coibion, Olivier, Yuriy Gorodnichenko, and Mauricio Ulate.** 2019. “The cyclical sensitivity in estimates of potential output.” *Brookings Papers on Economic Activity (forthcoming)*.
- Congdon, Tim.** 2008. “Two concepts of the output gap.” *World Economics*, 9(1): 147.
- Daly, Mary C, Bart Hobijn, Timothy Ni, et al.** 2013. “The path of wage growth and unemployment.” *FRBSF Economic Letter*, 20: 1–5.
- Dickens, William T, Lorenz Goette, Erica L Groshen, Steinar Holden, Julian Messina, Mark E Schweitzer, Jarkko Turunen, and Melanie E Ward.** 2007. “How wages change: micro evidence from the International Wage Flexibility Project.” *Journal of Economic Perspectives*, 21(2): 195–214.
- Dupraz, Stephane, Emi Nakamura, and Jon Steinsson.** 2017. “A plucking model of business cycles.” *Unpublished manuscript*.
- Fahr, Stephan, and F Smets.** 2010. “Downward wage rigidities in a monetary union with optimal monetary policy.” *The Scandinavian Journal of Economics*, 112(4): 812–840.
- Fatás, Antonio.** 2018. “Fiscal Policy, Potential Output, and the Shifting Goalposts.” *IMF Economic Review*, 1–19.
- Friedman, Milton.** 1968. “The role of monetary policy.” *American Economic Review*, 58-1.
- Kangur, Alvar, Koralai Kirabaeva, Jean-Marc Natal, and Simon Voigts.** 2019. “How Informative Are Real Time Output Gap Estimates in Europe?” *IMF Working Paper, forthcoming*.
- Kempkes, Gerhard.** 2012. “Cyclical adjustment in fiscal rules: Some evidence on real-time bias for EU-15 countries.”
- Marcellino, Massimiliano, and Alberto Musso.** 2011. “The reliability of real-time estimates of the euro area output gap.” *Economic Modelling*, 28(4): 1842–1856.
- Okun, Arthur M.** 1962. “Potential GNP Its Measurement and Significance.” *American Statistical Association, Proceedings of the Business and Economics Statistics Section*, 98-104.
- Orphanides, Athanasios, and Simon van Norden.** 2002. “The unreliability of output-gap estimates in real time.” *Review of economics and statistics*, 84(4): 569–583.

- Phelps, Edmund S.** 1967. “Phillips curves, expectations of inflation and optimal unemployment over time.” *Economica*, 254–281.
- Phillips, Alban W.** 1958. “The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957 1.” *Economica*, 25(100): 283–299.
- Pissarides, Christopher A.** 2000. *Equilibrium unemployment theory*. MIT press.
- Ravn, Morten O, and Harald Uhlig.** 2002. “On adjusting the Hodrick-Prescott filter for the frequency of observations.” *Review of economics and statistics*, 84(2): 371–376.
- Schmitt-Grohé, Stephanie, and Martin Uribe.** 2016. “Downward nominal wage rigidity, currency pegs, and involuntary unemployment.” *Journal of Political Economy*, 124(5): 1466–1514.
- Smets, Frank, and Rafael Wouters.** 2007. “Shocks and frictions in US business cycles: A Bayesian DSGE approach.” *American economic review*, 97(3): 586–606.
- Vetlov, Igor, Tibor Hlédik, Magnus Jonsson, Kucsera Henrik, and Massimiliano Pisani.** 2011. “Potential output in DSGE models.” *ECB Working Paper*, , (1351).